

## **NOVEL METHOD AND PHAGE FOR THE IDENTIFICATION OF NUCLEIC ACID SEQUENCES ENCODING MEMBERS OF A MULTIMERIC (POLY)PEPTIDE COMPLEX**

The present invention relates to methods for the identification of nucleic acid sequences encoding members of a multimeric (poly)peptide complex by screening for polyphage particles. Furthermore, the invention relates to products and uses thereof for the identification of nucleic acid sequences in accordance with the present invention.

Since its first conception by Ladner in 1988 (WO88/06630), the principle of displaying repertoires of proteins on the surface of phage has experienced a dramatic progress and has resulted in substantial achievements. Initially proposed as display of single-chain Fv (scFv) fragments, the method has been expanded to the display of bovine pancreatic trypsin inhibitor (BPTI) (WO90/02809), human growth hormone (WO92/09690), and of various other proteins including the display of multimeric proteins such as Fab fragments (WO91/17271; WO92/01047).

A Fab fragment consists of a light chain comprising a variable and a constant domain (VL-CL) non-covalently binding to a heavy chain comprising a variable and constant domain (VH-CH1). In Fab display one of the chains is fused to a phage coat protein, and thereby displayed on the phage surface, and the second is expressed in free form, and on contact of both chains, the Fab assembles on the phage surface.

Various formats have been developed to construct and screen Fab phage-display libraries. In its simplest form, just one repertoire, e. g. of heavy chains, is encoded on the phage or phagemid vector. A corresponding light chain, or a repertoire of light chains, is expressed separately. The Fab fragments assemble either inside a host cell, if the light chain is co-expressed from a plasmid, or outside the cell in the medium, if a collection of secreted phage particles each displaying a heavy chain is contacted with the light chain(s) expressed from a different host cell. By screening such Fab libraries, just the information about the heavy chain encoded on the phage or phagemid vector is retrievable, since that vector is packaged in the phage particle. By reverting the format and displaying a library of light chains, and assembling

Fab fragments by co-expressing or adding one or more of the heavy chains identified in the first round, corresponding light chain-heavy chain pairs can be identified.

To avoid that multi-step procedure, both repertoires may be cloned into one phage or phagemid vector, one chain expressible as a fusion with at least part of a phage coat protein, the second expressible in free form. After selection, the phage particle will contain the sequence information about both chains of the selected Fab fragments. The disadvantage of such a format is that the overall complexity of the library is limited by transformation efficiency. Therefore, the library size will usually not exceed  $10^{10}$  members.

For various applications, a library size of up to  $10^{14}$  would be advantageous. Therefore, methods of using site-specific recombination, either based on the Cre/lox system (WO92/20791) or on the att $\lambda$  system (WO 95/21914) have been proposed. Therein, two collection of vectors are sequentially introduced into host cells. By providing the appropriate recombination sites on the individual vectors, recombination between the vectors can be achieved by action of an appropriate recombinase or integrase, achieving a combinatorial library, the overall library size being the product of the sizes of the two individual collections. The disadvantages of the Cre/lox system are that the recombination event is not very efficient, it leads to different products and is reversible. The att $\lambda$  system leads to a defined product, however, it creates one very large plasmid which has a negative impact on the production of phages. Furthermore, the action of recombinase or integrase most likely leads to undesired recombination events.

Thus, the technical problem underlying the present invention is to develop a simple, reliable system which enables the simultaneous identification of members of a multimeric (poly)peptide complex, such as the identification of heavy and light chain of a Fab fragment, in phage display systems.

The solution to this technical problem is achieved by providing the embodiments characterized in the claims. Accordingly, the present invention allows to easily create and screen large libraries of multimeric (poly)peptide complexes for properties such as binding to a target, as in the case of screening Fab fragment libraries, or such as enzymatic activity, as in the case of libraries of multimeric enzymes. The technical approach of the present invention, i.e. the retrieval of information about two members of a multimeric (poly)peptide complex

encoded on two different vectors without requiring a recombination event, is neither provided nor suggested by the prior art.

Accordingly, the present invention relates to a method for identifying a combination of nucleic acid sequences encoding two members of a multimeric (poly)peptide complex with a predetermined property, said combination being contained in a combinatorial library of phage particles displaying a multitude of multimeric (poly)peptides complexes, said method being characterized by screening or selecting for polyphage particles that contain said combination.

Surprisingly, it has been achieved by the present invention that the phenomenon of polyphages can be used to co-package the genetic information of two or more members of multimeric (poly)peptide complexes in a phage display system. The occurrence of polyphage particles has been observed 30 years ago (Salivar et al., Virology 32 (1967) 41-51), where it was described that approximately 5% of a phage population form particles which are longer than unit length and which contain two or more copies of phage genomic DNA. They occur naturally when a newly forming phage coat encapsulates two or more single-stranded DNA molecules. In specific cases, it has been seen that co-packaging of phage and phagemids or single-stranded plasmid vectors takes place as well (Russel and Model, J. Virol. 63 (1989) 3284-3295). Despite of occasional scientific articles about the morphogenesis of polyphage particles, a practical application has never been discussed or even been mentioned. In WO92/20791 in example 26, a model experiment for a combinatorial Fab display library expressed from separate vectors is presented. However, there is only a screening process for either of the two vectors described. Thus, the prior art teaches away from screening for the simultaneous presence of two vectors in a polyphage particle.

In the context of the present invention, the term "multimeric (poly)peptide complex" refers to a situation where two or more (poly)peptide(s) or protein(s), the "members" of said multimeric complex, can interact to form a complex. The interaction between the individual members will usually be non-covalent, but may be covalent, when post-translational modification such as the formation of disulphide-bonds between any two members occurs. Examples for "multimeric (poly)peptide complexes" comprise structures such as fragments derived from immunoglobulins (e. g. Fv, disulphide-linked Fv (dsFv), Fab fragments), fragments derived from other members of the immunoglobulin superfamily (e.g.  $\alpha$ , $\beta$ -

heterodimer of the T-cell receptor), and fragments derived from homo-or heterodimeric receptors or enzymes. In phage display, one of said members is fused to at least part of a phage coat protein, whereby that member is displayed on, and assembly of the multimeric complex takes place at, the phage surface. A "combinatorial phage library" is produced by randomizing at least two members of said multimeric (poly)peptide complex at least partially on the genetic level to create two libraries of genetically diverse nucleic acid sequences in appropriate vectors, by combining the libraries in appropriate host cells and by achieving co-expression of said at least two libraries in a way that a library of phage particles is produced wherein each particle displays one of the possible combinations out of the two libraries.

By screening such a combinatorial phage library displaying multimeric (poly)peptide complexes for a predetermined property, a collection of phage particles will be identified. Partially, these particles will just contain the genetic information of one of the members of the multimeric complex. The inventive principle of the present invention is the screening step for polyphage particles containing the genetic information of a combination of library members.

Furthermore, the present invention relates to a method for identifying a combination of nucleic acid sequences encoding two members of a multimeric (poly)peptide complex with a predetermined property, said combination being contained in a combinatorial library of phage particles displaying a multitude of multimeric (poly)peptides complexes, comprising the steps of

- (a) providing a first library of recombinant vector molecules containing genetically diverse nucleic acid sequences comprising a variety of nucleic acid sequences, each encoding a fusion protein of a first member of a multimeric (poly)peptide complex fused to at least part of a phage coat protein, said fusion protein thereby being able to be directed to, and displayed at, the phage surface, wherein said vector molecules are able to be packaged in a phage particle and carry or encode a first selectable and/or screenable property;
- (b) providing a second library of recombinant vector molecules containing genetically diverse nucleic acid sequences comprising a variety of nucleic acid sequences, each encoding a second member of a multimeric (poly)peptide complex, wherein the vector molecules of said second library are able to be packaged in a phage particle and carry or encode a second selectable and/or screenable property different from said first property;

- (c) optionally, providing nucleic acid sequences encoding further members of a multimeric (poly)peptide complex;
- (d) expressing members of said libraries of recombinant vectors mentioned in steps (a), (b), and optionally nucleic acid sequences mentioned in step (c), in appropriate host cells under appropriate conditions, so that a combinatorial library of phage particles each displaying a multimeric (poly)peptide complex is produced;
- (e) identifying in said library of phage particles a collection of phages displaying multimeric (poly)peptide complexes having said predetermined property;
- (f) identifying in said collection polyphage particles simultaneously containing recombinant vector molecules encoding a first and a second member of said multimeric (poly)peptide complex by screening or selecting for the simultaneous presence or generation of said first and second selectable and/or screenable property;
- (g) optionally, carrying out further screening and/or selection steps or repeating steps (a) to (f);
- (h) identifying said combination of nucleic acid sequences.

Optionally, further members of said multimeric complex may be provided in the case of ternary, quaternary or higher (poly)peptide complexes. These further members may, for example, be co-expressed from one of the phage or phagemid vectors or from a separate vector such as a plasmid. Even libraries of such further members could be employed in which case further screenable or selectable properties would have to be introduced on the corresponding vectors. Alternatively, such further libraries could be contained in said first or second libraries of recombinant vector molecules. In another option, further screening and/or selection steps or a repetition of the individual steps can be carried out, to optimize the result of obtaining and identifying said nucleic acid sequences.

Furthermore, the present invention relates to a method for identifying a combination of nucleic acid sequences encoding two members of a multimeric (poly)peptide complex with a predetermined property, said combination being contained in a combinatorial library of phage particles displaying a multitude of multimeric (poly)peptides complexes, comprising the steps of

- (a) expressing in appropriate host cells under appropriate conditions
  - (aa) genetically diverse nucleic acid sequences contained in a first library of recombinant vector molecules, said nucleic acid sequences comprising a variety



of nucleic acid sequences, each encoding a fusion protein of a first member of a multimeric (poly)peptide complex fused to at least part of a phage coat protein, said fusion protein thereby being able to be directed to and displayed at the phage surface, wherein said vector molecules are able to be packaged in a phage particle and carry or encode a first selectable and/or screenable property;

(ab) genetically diverse nucleic acid sequences contained in a second library of recombinant vector molecules, said nucleic acid sequences comprising a variety of nucleic acid sequences, each encoding a second member of a multimeric (poly)peptide complex, wherein the vector molecules are able to be packaged in a phage particle and carry or encode a second selectable and/or screenable property different from said first property;

(ac) optionally, nucleic acid sequences encoding further members of a multimeric (poly)peptide complex,

so that a combinatorial library of phage particles each displaying a multimeric (poly)peptide complex is produced;

(b) identifying in said library of phage particles a collection of phages displaying multimeric (poly)peptide complexes having said predetermined property;

(c) identifying in said collection polyphage particles simultaneously containing recombinant vector molecules encoding a first and a second member of said multimeric (poly)peptide complex by screening or selecting for the simultaneous presence or generation of said first and second selectable and/or screenable property;

(d) optionally, carrying out further screening and/or selection steps or repeating steps (a) to (c);

(e) identifying said combination of nucleic acid sequences.

In a preferred embodiment of the method of the present invention, the vectors of said first and said second library are a combination of a phage vector and a phagemid vector.

In a further preferred embodiment of the method of the present invention, the vectors of said first and said second library are a combination of two phagemid vectors, said appropriate conditions comprising complementation of phage genes by a helper phage.

In a most preferred embodiment of the method of the present invention said two phagemid vectors are compatible.

The term "compatibility" refers to a property of two phagemids to be able to coexist in a host cell. Incompatibility is connected to the presence of incompatible plasmid origins of replication belonging to the same incompatibility group. An example for compatible plasmid origins of replication is the high-copy number origin ColE1 and the low-copy number origin p15A.

Therefore, in a further preferred embodiment of the method of the present invention, said two phagemid vectors comprise a ColE1 and a p15A plasmid origin of replication.

In a most preferred embodiment of the method of the present invention, said two phagemid vectors comprise a ColE1 and a mutated ColE1 origin.

It could be shown, that two phagemids both having a ColE1-derived plasmid origin of replication can coexist in a cell as long as one of the ColE1 origins carries a mutation.

Particularly preferred is a method, wherein said vectors and/or said helper phage comprise different phage origins of replication.

Most preferred is an embodiment of the method of the present invention, wherein said phage vector, said phagemid vector(s) and/or said helper phage are interference resistant.

The term "interference" refers to a property that phagemids inhibit the production of progeny phage particles by interfering with the replication of the DNA of the phage. "Interference resistance" is a property which overcomes this problem. It has been found that mutations in the intergenic region and/or in gene II contribute to interference resistance (Enea and Zinder, Virology 122 (1982), 222-226; Russel et al., Gene 45 (1986) 333-338). It was identified that phages called IR1 and IR2 (Enea and Zinder, Virology 122 (1982), 222-226), and mutants derived therefrom such as R176 (Russel and Model, J. Bacteriol. 154 (1983) 1064-1076), R382, R407 and R408 (Russel et al., Gene 45 (1986) 333-338) and R383 (Russel and Model, J. Virol. 63 (1989) 3284-3295) are interference resistant by carrying mutations in the untranslated region upstream of gene II and in the gene II coding region.

Therefore, in a preferred embodiment of the method of the present invention, said phage vector, said phagemid vector(s) and/or said helper phage have mutations in the phage intergenic region(s), preferably in positions corresponding to position 5986 of f1, and/or in gene II, preferably in positions corresponding to position 143 of f1.

In a most preferred embodiment said phage vector, said phagemid vector(s) and/or said helper phage are, or are derived from, IR1 mutants such as R176, R382, R383, R407, R408, or from IR2 mutants.

In a further embodiment of the method of the invention, said vectors and/or said helper phage comprise hybrid nucleic acid sequences of f1, fd, and/or M13 derived sequences.

In the context of the present invention, the term "hybrid nucleic sequences" refers to vector elements which comprise sequences originating from different phage(mid) vectors.

Surprisingly, it has been found that a vector constructed combining a part derived from fd phage and a second part derived from R408, a derivative of f1 phages, is interference resistant and additionally, gives predominantly polyphage particles.

Therefore, a most preferred embodiment of the method of the present invention relates to a vector which is, or is derived from, fpep3\_1B-IR3seq with the sequence listed in Figure 4.

In a yet further preferred embodiment of the method according to the present invention, said derivative is a phage comprising essentially the phage origin or replication from fpep3\_1B-IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

The invention relates in an additional preferred embodiment to a method, wherein said derivative is a phagemid comprising essentially the phage origin or replication from fpep3\_1B-IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

The invention relates in a further preferred embodiment to a method, wherein said derivative is a helper phage comprising essentially the phage origin or replication from fpep3\_1B-



IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

Most preferred is an embodiment of the method of the invention, wherein said derivatives comprise the combined fd/f1 origin including the mutation G5737>A (2976 in fpep3\_1B-IR3seq), and/or the mutations G343>A (3989) in gII, and G601>T (4247) in gII/X.

The formation of polyphage particles has been examined in more detail by different groups. It was found that amber mutations in genes VII and IX lead to the amplified production of infectious polyphage particles (Lopez and Webster, Virology 127 (1983) 177-193). A couple of mutants in gene VII (R68, R100) and in gene IX (N18) were identified and further characterized.

Accordingly, in a preferred embodiment of the method of the present invention, the gene VII contained in any of said vectors contains an amber mutation, and most preferably, said mutation is identical to those found in phage vectors R68 or R100.

Further preferred is an embodiment, wherein the gene IX contained in any of said vectors contains an amber mutation, and most preferably said mutation is identical to that found in phage vector N18.

Several phage coat proteins have been used in displaying foreign proteins including the gene III protein (gIIIp), gVIp, and gVIIIp.

In a preferred embodiment of the method of the present invention, said phage coat protein is gIIIp or gVIIIp.

In a particularly preferred embodiment of the method of the present invention, said phage particles are infectious by having a full-length copy of gIIIp.

The gIIIp is a protein comprising three domains. The C-terminal domain is responsible for membrane insertion, the two N-terminal domains are responsible for binding to the F pilus of *E. coli* (N2) and for the infection process (N1).

In a most preferred embodiment of the method of the invention, said phage particles are non-infectious by having no full-length copy of gIIIp, said fusion protein being formed with a truncated version of gIIIp, wherein the infectivity can be restored by interaction of the

displayed multimeric (poly)peptide complexes with a corresponding partner coupled to an infectivity-mediating particle.

In the context of the present invention, the term "infectivity-mediating particle" (IMP) refers to a construct comprising either the N1 domain or the N1-N2 domain. On interaction with a non-infectious phage lacking said domains, infectivity of the phage particles can be restored. The interaction between the non-infectious phage and the IMP can be mediated by a ligand fused to the IMP, which can bind to a partner displayed on the phage. By screening a non-infectious phage display library against a target ligand-IMP construct, restoration of infectivity can be used to select target-binding library members.

In a further preferred embodiment of the method of the invention, said truncated gIIIp comprises the C-terminal domain of gIIIp.

In a yet preferred embodiment of the method of the invention, said truncated gIIIp is derived from phage fCA55.

In addition to the work by Lopey and Webster cited above, Crissman and Smith (Virology 132 (1984) 445-455) could show, that the phage fCA55 which has a large deletion in gene III removing the N-terminal domains and a large part of the C-terminal domain leads exclusively to the formation of polyphages.

Particularly preferred is an embodiment of the method of the invention, wherein said predetermined property is binding to a target.

In a preferred embodiment of the method of the invention, said multimeric (poly)peptide complex is a fragment of an immunoglobulin superfamily member.

In a most preferred embodiment of the method of the invention, said multimeric (poly)peptide complex is a fragment of an immunoglobulin.

In a further most preferred embodiment of the method of the invention, said fragment is an Fv, dsFv or Fab fragment.

An additional preferred embodiment of the present invention relates to a method, wherein said predetermined property is the activity to perform or to catalyze a reaction.

In a preferred embodiment of the method of the invention, said multimeric (poly)peptide complex is an enzyme.

In a most preferred embodiment of the method of the invention, said multimeric (poly)peptide complex is a fragment of a catalytic antibody.

In a further most preferred embodiment of the method of the invention, said fragment is an Fv, dsFv or Fab fragment.

An additional preferred embodiment of the invention relates to a method, wherein selectable and/or screenable property is the transactivation of transcription of a reporter gene such as beta-galactosidase, alkaline phosphatase or nutritional markers such as his3 and leu, or resistance genes giving resistance to an antibiotic such as ampicillin, chloramphenicol, kanamycin, zeocin, neomycin, tetracycline or streptomycin.

In a most preferred embodiment of the method of the invention, said generation of said first and second screenable and/or selectable property is achieved after infection of appropriate host cells by said collection of phage particles.

Particularly preferred is a method, wherein said identification of said nucleic acid sequences is effected by sequencing.

Further preferred is a method, wherein said host cells are E.coli XL-1 Blue, K91 or derivatives, TG1, XL1kann or TOP10F.

An additional preferred embodiment of the invention relates to a polyphage particle which

(a) contains

(i) a first recombinant vector molecule that comprises a nucleic acid sequence, which encodes a fusion protein of a first member of a multimeric (poly)peptide complex

fused to at least part of a phage coat protein, and that carries or encodes a first selectable and/or screenable property, and

(ii) a second recombinant vector molecule that comprises a nucleic acid sequence, which encodes a second member of a multimeric (poly)peptide complex, and that carries or encodes a second selectable and/or screenable property different from said first property;

and (b) displays said multimeric (poly)peptide complex at its surface.

A most preferred embodiment of the invention relates to a polyphage particle, wherein said phage coat protein is the gIIIp.

A further preferred embodiment of the present invention relates to a polyphage particle which is infectious by having a full-length copy of gIIIp present, either in said fusion protein, or in an additional wild-type copy.

Additionally, the invention relates to a polyphage particle which is non-infectious by having no full-length copy of gIIIp, said fusion protein being formed with a truncated version of gIIIp, wherein the infectivity can be restored by interaction of the displayed multimeric (poly)peptide complex with a corresponding partner coupled to an infectivity-mediating particle.

Most preferably, the invention relates to the phage vector fpep3\_1B-IR3seq with the sequence listed in Figure 4.

Additionally preferred, the invention relates to a phage vector derived from phage vector fpep3\_1B-IR3seq comprising essentially the phage origin or replication from fpep3\_1B-IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

Further preferred is an embodiment of the invention, which relates to a phagemid vector derived from phage vector fpep3\_1B-IR3seq comprising essentially the phage origin of replication from fpep3\_1B-IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

Preferably, the invention relates to a helper phage vector derived from phage vector fpep3\_1B-IR3seq comprising essentially the phage origin of replication from fpep3\_1B-IR3seq, the gene II from fpep3\_1B-IR3seq, or a combination of said phage origin of replication and said gene II.

Additionally preferred is an embodiment, said derivatives comprise the combined fd/f1 origin including the mutation G5737>A (2976 in fpep3\_1B-IR3seq), and/or the mutations G343>A (3989) in gII, and G601>T (4247) in gII/X.

Further preferred is the use of any of the vectors according to the present invention in the generation of polyphage particles containing a combination of at least two different vectors.

Most preferred is the use of vectors of the invention, wherein said combination of different vectors comprises nucleic acid sequences encoding members of a multimeric (poly)peptide complex.

Further preferred in the present invention is the use of vectors, wherein said combination of different vectors comprises nucleic acid sequences encoding interacting (poly)peptides/proteins.

### **Legends to Figures:**

**Figure 1:** General description of the polyphage principle for the display of a Fab library: e.g. library 1: library of VL chains; library 2: VH chains; both libraries on compatible phagemids; in a: libraries are transformed into host cells; in b: library 1 is rescued by a helper phage; in c: libraries are combined by infection; in d: co-expression of heavy and light chains; in e: rescue by helper phages, production of phage particles, assembly of Fab on phage, selection for target; note 1: A certain fraction of the phage particles will be normal unit-length particles containing just one of the two genomes (not shown in Figure 1). Furthermore, polyphage does not discriminate which genomes to package. Therefore, the combinations shown in Figure 1 can arise. To select for correctly



packaged genomes, the subsequent steps are required; in f: infect host cells; in g: select for ability to confer resistance to two antibiotics to infected cells; note 2: only phage that satisfy condition according to g) represent polyphage particles which contain the correct combination of heavy and light chain of binding Fabs (Hetero-polyphage). Unit-length phage as well as polyphage carrying two identical genomes will confer only resistance to one antibiotics.

- Figure 2: Functional map and sequence of phage vector fhag1A
- Figure 3: Functional map and sequence of phage vector fjun\_1B
- Figure 4: Functional map and sequence of phage vector fpep3\_1B-IR3seq
- Figure 5: Compatibility of various phage and phagemid vectors: co-transformation of different vector pairs and growth in liquid culture (can/amp selection):  
 A. fjun\_1B-R408-IR/pIG10\_pep10; B. fjun\_1B/pIG10\_pep10 (only 1 colonie);  
 C. fpep3\_1B-IR3/pIG10\_pep10; D. fjun\_1B-R408-IR/pOK1Djun; E. fjun\_1B/pOK1Djun: no growth; F. fpep3\_1B-IR3/pOK1Djun;  
 a. fjun\_1B; b. fjun\_1B-R408-IR; c. fpep3\_1B-IR3; d. pIG10\_pep10; e. pOK1Djun
- Figure 6: co-transformation of positive (pep3/p75ICD combination, lane 9) and negative (jun/p75ICD, lane 10) pairs; lane 1 to 8: SIP transductants
- Figure 7: Sensitivity of SIP hetero-polyphage system for selection in solution: #SIP hetero-polyphage transductants, transducing units (t.u.)/ml, produced by co-cultures of co-transformants as in Figure 6 mixed at the indicated ratios.
- Figure 8: PCR to identify phage vector(s) present in SIP polyphage transductants: lane 1 to 6: SIP polyphage transductants; lane A: fpep3\_1B-IR3/pIG10.3-IMPp75 co-transformant; lane B: fjun\_1B-IR3/pIG10.3-IMPp75 co-transformant
- Figure 9: IR Phage and Phagemid are Co-packaged into Polyphages: 1:  $\Delta$ gIII phage + gIII plasmid; 2: IR phage+ phagemid
- Figure 10: SIP Information is Co-transduced by Polyphages: a: IMPp75 on phage vector; b: pep10-gIII-CT fusion on phage vector; c: IMPp75 on phagemid vector; d: pep10-gIII-CT fusion on phagemid vector

The examples illustrate the invention

### **Example 1: Selection for polyphage transductants**

In WO92/01047, page 83, a model experiment for a two-vector system is described which uses a phage vector (fd-CAT2-IV) encoding a light chain and a phagemid vector (pHEN1-III) encoding a heavy chain. The phagemid, grown in *E. coli* HB2151, was rescued with fd-CAT2-IV phage, and functional phage(mid)s produced. By infecting TG1 cells and plating on tetracycline (to select for fd-CAT) and ampicillin (to select for pHEN1), the ratio of phage and phagemid being packaged was determined.

By repeating this experiment, but plating on TYE plates with both antibiotics, polyphage transductants transducing both resistances simultaneously can be selected, and the genetic information contained on the phage and phagemid vector can be retrieved.

By replacing the single light and heavy chain in the constructs mentioned above by corresponding repertoires, a library of Fab-displaying phage particles can be produced. By screening that library against an immobilized target, a collection of phage particles can be identified. Polyphage particles contained in that collection can be identified by transducing both resistances as described above.

### **Example 2: Generation and use of an interference-resistant filamentous phage to co-package the genetic information of co-displayed interacting proteins**

#### **Introduction**

The physical connection of randomly combined genetic information is of vital importance in processes such as interactive screening of two libraries of expressed protein members or for co-expression and co-display of protein pairs which are dependent on the interaction with each other for proper function.

#### **2.1.: Construction of a interference resistant filamentous phage:**

##### **2.1.1.: Construction of fjun\_1B:**

##### **- fhag1A (see Figure 2)**

- a. The phage vector f17/9-hag (Krebber *et al.*, 1995, *FEBS Letters* 377, 227-231) is digested with EcoRV and XmnI. The 1.1 kb fragment containing the anti-HAG Ab gene is isolated

by agarose gel electrophoresis and purified with a Qiagen gel extraction kit. This fragment is ligated into a pre-digested pIG10.3 vector (EcoRV-XmnI). Ligated DNA is transformed into DH5a cells and positive clones are verified by restriction analysis. The recombinant clone is called **pIGhag1A**. All cloning described above and subsequently are according to standard protocols (Sambrook *et al.*, 1989, *Molecular Cloning: a Laboratory Manual*, 2<sup>nd</sup> ed.)

- b. The vector f17/9-hag (Krebber *et al.*, 1995) is digested with EcoRV and StuI. The 7.9 kb fragment is isolated and self-ligated to form the vector **fhag2**.
- c. The chloramphenicol resistance gene (CAT) assembled *via* assembly PCR (Ge and Rudolph, *BioTechniques* 22 (1997) 28-29) using the template pACYC (Cardoso and Schwarz, *J. Appl. Bacteriol.* 72 (1992) 289-293) is amplified by the polymerase chain reaction (PCR) with the primers:  
CAT\_BspEI(for): 5' GAATGCTCATCCGGAGTTC  
CAT\_Bsu36I(rev): 5' TTTCAGTGGCCTCAGGCTAGCACCGGCGTTTAAG
- d. The PCR is done following standard protocols (Sambrook *et al.*, 1989). The amplified product is digested with BspEI and Bsu36I then ligated into pre-digested fhag2 vector (BspEI-Bsu36I; 7.2 kb fragment) to form **fhag2C**.
- e. The vector fhag2C is digested with EcoRI and the ends made blunt by filling-in with Klenow fragment. The flushed vector is self-ligated to form vector **fhag2CdelEcoRI**.
- f. pIGhag1A is digested with XbaI and HindIII. The 1.3 kb fragment containing the anti-HAG gene fused with the C-terminal domain of filamentous phage pIII protein is isolated and ligated with a pre-digested fhag2CdelEcoRI phage vector (XbaI-HindIII; 6.4 kb) to create the vector **fhag1A**.

**- fjun\_1B (see Figure 3)**

- a. The DNA encoding the C-terminal domain including the long linker separating it from the amino terminal domain of the filamentous phage pIII (gIII short) is amplified by PCR using pOK1 (Gramatikoff *et al.*, *Nucleic Acids Res.* 22 (1994) 5761-5762) as template with the primers:  
gIII short(for): 5'GCTTCCGGAGAATTCAATGCTGGCGGCGGCTCT3'  
gIII short(rev): 5'CCCCCCCCAAGCTTATCAAGACTCCTTATTACG3'
- b. The PCR is done following standard protocols (Sambrook *et al.*, 1989). The amplified product is digested with EcoRI and HindIII, then ligated into pre-digested fhag1A vector (EcoRI-HindIII) to form the vector **fjun\_1B**.

### 2.1.2.: Construction of fjun\_1B-R408IR:

In order to introduce mutations which have been described to confer an interference resistance phenotype (Enea and Zinder, Virology 122 (1982), 222-226) into the non-interference resistant fd phage vector fjun\_1B (see Fig.3), a 1.7 kb fragment of helper phage R408 (Stratagene) comprising the region between the unique restriction sites *DraIII* and *BsrGI* was PCR amplified by assembly PCR. Subfragments of the 1.7 kb *DraIII*/*BsrGI* fragment were amplified from the f1 phage R408 template DNA with primer combinations FR604/FR605 and FR606/FR607 to introduce via the partially complementary primers FR605 and FR606 an additional *gII* mutation found to be present in the recipient construct fjun\_1B. Resulting PCR fragments were gel-purified and combined to serve as template in an subsequent assembly PCR with primers FR604 and FR607. PCR conditions were standard, with approx. 25 ng template, 10 pmole of each primer, 250 pmole of each dNTP, 2 mM Mg, 2.5 U Pfu DNA polymerase (Stratagene). Amplification was done for 30 cycles, with 1 min denaturation at 94 C, 1 min annealing at 50°C, 1 min extension at 72°C. The correct-sized 1.7 kb assembly PCR product was gel-purified, digested with *DraIII* and *BsrGI* and cloned into *DraIII*/*BsrGI*-digested fjun\_1B, generating fjun\_1B-R408IR.

Primers:      FR604 5' GTTCACGTAGTGGGCCATCG 3'  
                  FR605 5' TGAGAGGTCTAAAAAGGCTATCAGG 3'  
                  FR606 5' TAGCCTTTTGTAGACCTCTCAAAAATAG 3'  
                  FR607 5' CGGTGTACAGACCAGGCGC 3'

### 2.2.: Proof of principle experiments

Despite of the absence of the two originally associated IR mutations, the hybrid phage vector fjun\_1B-R408IR (carrying the chloramphenicol acetyltransferase conferring chloramphenicol resistance) could be co-transformed with a phagemid (pOK1deltajun, carrying the beta-lactamase gene conferring ampicillin resistance) containing a phage origin of replication. More importantly, fjun\_1B-R408IR could stably co-exist with the phagemid pOK1deltajun, and the phagemid was efficiently co-packaged together with the fjun\_1B-R408IR phage genome into polyphage particles. Titers of polyphages, simultaneously

transducing chloramphenicol and ampicillin resistance, reached  $6 \times 10^8$  transducing units (t.u.)/ml of overnight bacterial culture K91 plating cells, a number almost equivalent to a titer of  $10^9$ /ml seen after selection on chloramphenicol only. Selection of the K91 transductants on ampicillin only gave a titer of  $5 \times 10^9$ /ml. These titers indicated that more than 50 % of all phages containing fjun\_1B-R408IR also contained the phagemid pOK1deltajun, thus representing polyphages. This high ratio of polyphages was confirmed by restriction analysis of transductants which had been selected on chloramphenicol only. More than 50 % of these clones also contained the phagemid in addition to the fjun\_1B-R408IR phage genome. fjun\_1B-R408IR was isolated in pure form from an individual transductant, which contained only this phage. The construct fjun\_1B-R408IR was used with pOK1deltajun for co-transformation of DH5 $\alpha$  cells, in order to produce selectively-infective phages (SIP) via fos-jun leucine zipper interaction (which non-covalently restores wt gIII function). Stable, double-resistant co-transformants were obtained with this combination and individual clones were grown overnight in the presence of cam/amp. The culture supernatant of these clones was filtered through a 45  $\mu$ M membrane filter and used to infect exponentially-growing F+ bacteria (K91 strain) for 20 min at 37 C. To test for the presence of infective SIP polyphages the cells were plated on LB agar plates containing cam and amp and plates were incubated at 37 C overnight. Approx. 500 to 1000 transforming units (t.u.)/ml resulting in double-resistant transductants were obtained from individual co-transformants. DNA of those transductants was analyzed by restriction analysis which showed that 95 % (15/16 clones) of the clones had the correct pattern expected for fjun\_1B-R408IR and pOK1deltajun. Supernatants of several polyphage transductants were tested for persistent SIP phage production by re-infection of K91 cells. This confirmed that polyphage transductants continued to produce infective SIP phages and restriction analysis of the resulting 2<sup>nd</sup> round polyphage transductants showed that 44 % (14/32 clones) contained the correct vector combination. The rest of the clones contained the correct pOK1deltajun phagemid plus a recombined phage vector with a restored wt gIII, indicating an increase in recombination frequency when both vectors are propagated in the rec+ strain K91 (compared to the rec- strain DH5 $\alpha$  used for co-transformation of IR phage and phagemid). To test other protein-protein interactions which give a higher titer of infective SIP phages and to verify the presence of hetero-polyphages (co-packaging of phage and phagemid instead of co-infection by monophages or homo-polyphages) , two peptide ligands (previously selected by SIP, WO97/32017) which bind to the p75 rat



neurotrophin receptor (Chao et al., Science 232 (1986) 518-521) intracellular domain (p75ICD) were cloned as N-terminal gIIIc fusions in fjun\_1B-R408IR (replacing jun) and the phagemid pIG10.3, leading to constructs fpep3\_1B-IR3seq and pIG10.3-pep10 (WO97/32017), respectively, which contain the peptide pep3: 5'-TGTATTGTTTATCATGCTCATTATCTTGTTGCTAAGTGT-3' encoding the amino acid sequence (CysIleValTyrHisAlaHisTyrLeuValAlaLysCys) instead of the jun sequence. Sequencing of the respective parts of the transferred R408 fragment in fpep3\_1B-IR3seq revealed that neither of the two IR mutations (the G5986>A mutation from complementation group I in the gII 5' non-translated region, which should be found at position 3225 in fpep3\_1B-IR3seq, and the C143>T mutation (3789 in fpep3\_1B-IR3seq) from complementation group II leading to a Thr>Ile amino acid exchange in gII) were found to be present. However; the gII mutation G6090>T (3329 in fpep3\_1B-IR3seq), leading to a Leu>Val exchange, introduced by assembly PCR was present. Furthermore, three additional mutations compared to an f1 phage could be identified: G5737>A (2976 in fpep3\_1B-IR3seq) in the phage origin of replication, G343>A (3989) in gII, and G601>T (4247) in gII/X.

The functional map and the sequence of fpep3\_1B-IR3seq are given in Figure 4. This sequence was double-checked several times. It could be shown that differences in the sequence of fpep3\_1B-IR3seq compared to published sequence data could be explained by mutations already present in the starting constructs used for cloning fjun\_1B-R408IR and fpep3\_1B-IR3seq.

Co-transformation experiments (**Fig. 5**) using combinations of pIG10.3 or pOK1 phagemids (both with f1 oris) with fjun\_1B ("wt" fd phage), fjun\_1B-R408-IR (containing the DraIII/BsrGI fragment from R408) or fpep3\_1B-IR3 (containing the DraIII/BsrGI fragment from R408 and the PCR mutation) revealed that the PCR mutation is not necessary for the IR phenotype, at least judged by the ability to be co-transformable with a phagemid and the ability of individual co-transformants to grow in liquid culture (cam/amp selection).

Additionally, the interacting protein partner p75ICD was cloned as a C-terminal fusion to the infectivity-mediating domains (N1-N2) of gIII (infectivity-mediating particle (IMP) fusion) resulting in constructs fIMPp75-IR3 and pIG10.3-IMPp75.

The IR phage was tested with the SIP pairing fpep3\_1B-IR3seq3/ pIG10.3-IMPp75 (which gives a higher titer than fos/jun SIP) in the presence of the negative control combination fjun\_1B-IR3seq3/ pIG10.3-IMPp75 (**Fig. 6**). A SIP hetero-polyphage titer of  $1.5 \times 10^5$ /ml (cam/amp-resistant transductants) was achieved with fpep3\_1B-IR3seq3/ pIG10.3-IMPp75. To test SIP sensitivity in a model library vs. library setting, co-transformants of fpep3\_1B-IR3seq3/ pIG10.3-IMPp75 were diluted in an excess fjun\_1B-IR3/ pIG10.3-IMPp75 and the supernatant of the bacterial co-culture was assayed for SIP hetero-polyphages. This showed that down to a dilution of  $10^{-5}$  to  $10^{-6}$  can be recovered (**Fig. 7**).

To prove that only the correct phage vector is present in SIP polyphage transductants, DNA of positive (fpep3\_1B-IR3seq3/ pIG10.3-IMPp75) and negative (fjun\_1B-IR3/ pIG10.3-IMPp75) control co-transformants, as well as DNA from the SIP polyphage transductants derived from SIP phages produced by the mix of positive and negative control bacteria was analyzed by PCR (**Fig. 8**). Primers FR614 (5'-GCTCTAGATAACGAGGGC-3') and FR627 (5'-CGCAAGCTTAAGACTCCTTATTACGC-3') amplify the phage region from the start of ompA to the end of gIII. PCR products derived from fpep3\_1B-IR3seq3 and fjun\_1B-IR3 can be discriminated by size. Gel analysis of the above samples verified that only the expected fpep3\_1B-IR3seq3 phage was present in SIP polyphage transductants (6 analyzed).

To physically demonstrate the existence of hetero-polyphages (which have phage and phagemid co-packaged) when using the IR phage vector, phages produced by co-transformants of fIR3/pIG10.3-IMPp75 and as a control fjun\_1B/JB61 ("wt" phage plus complementing gIII plasmid) were separated on an agarose gel (**Fig. 9**). This showed that the fIR3/pIG10.3-IMPp75 combination produced substantially more slower migrating (thus bigger) phages than the fjun\_1B/JB61 control combination. The ratio was almost inversed. Elution of phages from various regions of the gel and subsequent titering of the eluate on plating cells showed that the upper gel region contained a significant portion of double resistance-transducing phages which thus can be regarded as hetero-polyphages.

The pairs fpep3\_1B-IR3 and pIG10.3-IMPp75 as well as fIMPp75-IR3 and pIG10.3-pep10 were co-transformed into DH5 $\alpha$ , individual cam/amp resistant clones were grown and the culture supernatant was tested on K91 cells for SIP phage production (Fig. 10). The combinations fpep3\_1B-IR3/pIG10.3-IMPp75 and fIMPp75-IR3/pIG10.3-pep10 gave a titer of  $1.5 \times 10^5$  t.u./ml and  $5 \times 10^3$  t.u./ml, respectively when assayed for cam/amp-resistant transductants. The titer for each combination when assayed on LB cam was nearly the same as when assayed on LB cam/amp. This demonstrated efficient co-packaging of phage and phagemid DNA to almost 100 %, as seen before with the initial fjun\_1B-R408IR and pOK1deltajun combination. To proof the existence of polyphages which individually co-transduce phage and phagemid DNA simultaneously, and to rule out the possibility of transduction of the two resistance markers by independent (and thus random) co-infection by two different phages which have only phage or phagemid packaged, a statistical test was performed. Defined, identical aliquots of bacterial culture supernatants of an individual co-transformant representing each of the two SIP vector combinations described above (fpep3\_1B-IR3/pIG10.3-IMPp75 and fIMPp75-IR3/pIG10.3-pep10) were either used individually to infect K91 cells followed by selection on LB cam and LB amp plates, or the same supernatant aliquots from the two vector combinations were mixed before infection of K91 cells and selection on LB cam/amp. 117 cam-resistant, 328 amp-resistant and 141 cam/amp-resistant transforming units were present in the supernatant aliquot from the fIMPp75-IR3/pIG10.3-pep10 combination and 40 cam-resistant, 30 amp-resistant and 23 cam/amp-resistant transforming units were present in the supernatant aliquot from the fpep3\_1B-IR3/pIG10.3-IMPp75 combination. The mix of both supernatant aliquots contained 166 cam-resistant and 162 cam/amp-resistant transforming units, exactly corresponding to the expected numbers which would be obtained by adding up the transducing units of the two individual aliquots. 48 cam/amp-resistant transductant colonies were picked from the plate where the mix of the two individual aliquots was used for infection and were analyzed by restriction digest. This showed that only the correct, SIP phage-producing vector combination (5 clones containing the fpep3\_1B-IR3/pIG10.3-IMPp75 and 43 clones containing the fIMPp75-IR3/pIG10.3-pep10 combination; this represents a ratio of the two input vector combinations in the analyzed transductants of 1 : 8.6 (fpep3\_1B-IR3/pIG10.3-IMPp75 : fIMPp75-IR3/pIG10.3-pep10), which is very similar to the 1 : 6.1 (fpep3\_1B-IR3/pIG10.3-IMPp75 : fIMPp75-IR3/pIG10.3-pep10) ratio of double-resistant input phages in this experiment) occurred in all analyzed transductants,

verifying the presence of hetero-polyphages by ruling out the possibility of random co-infection and thus incorrect, random combination by two out of four possible monophage and/or homo-polyphage populations (fpep3\_1B-IR3, pIG10.3-IMPp75, fIMPp75-IR3 and pIG10.3-pep10) each containing only one type of vector (phage or phagemid). Statistically, co-infection of the same bacterium by two separate phages was practically already excluded by the small numbers of infective phages containing at least one resistance marker (166 cam-resistant and 358 amp-resistant phages) which were used in the above experiment. Co-infection of the same bacterium (of a total of  $10^7$  bacteria) by one of the 166 cam-resistant phages and one of the 358 amp-resistant phages has a probability of  $6 \times 10^{-10}$ . Moreover, in this scenario incorrect combinations of individual phage and phagemid vectors (e.g. fpep3\_1B-IR3/ pIG10.3-pep10 and fIMPp75-IR3/ pIG10.3-IMPp75) would be possible. The fact that only the correct vector combinations were found in all 48 transductants analyzed from this experiment further proved that co-transduction by hetero-polyphage and not random co-infection by homo-polyphage or monophage was the mechanism by which double-resistance was transduced.

### **2.3.: Construction of a phage-display system for Fab display**

The constructs described in 3.2. can easily be modified to achieve the display of Fabs or a Fab library. In fpep3\_1B-IR3seq, the jun part can be replaced by a VL-CL light chain repertoire having the appropriate 3'- and 5'-restriction sites similarly as described for pep\_3 to construct fVL\_1B-R408IR. In pIG10.3-IMPp75, the IMPp75 construct can be replaced by a repertoire of VH-CH1 heavy chains. After co-transformation of both repertoires into host cells and expression, a library of phage particles displaying Fab fragments is produced. Since fpep3\_1B-IR3seq was set up for a SIP experiment by having just the C-terminal domain of gIII, the corresponding Fab-displaying phage particles are non-infectious. By adding a target molecule fused to an infectivity-mediating particle (N1-N2 domain of gIIIp), phages displaying target-binding Fab fragments can be selected by infecting host cells.

By replacing the truncated gIII part described above by a full-length copy of gIII, a Fab-display library of infectious phage particles is obtained, which can be screened against immobilized targets. Binding phages can be eluted and used to infect host cells.

By selecting for transductants conferring cam/amp-resistance to their host cells, polyphage infections can be selected in both cases. Thereby the information about both chains of the selected Fab fragments can be retrieved.